Genome-enabled modeling of microbial biogeochemistry using a trait-based approach. Does increasing metabolic complexity increase predictive capabilities?

E. KING[1], S. MOLINS[1], U. KARAOZ[1], N.J. BOUSKILL[1], K. ANANTHARAMAN[2], H.R. BELLER[1], T. N. M. JEWELL[1], K.H. WILLIAMS[1], J.F. BANFIELD[1,2], C.I. STEEFEL[1], SS. HUBBARD[1], E.L. BRODIE[1,2*]

[1]Lawrence Berkeley National Laboratory, [2]University of California, Berkeley
*elbrodie@lbl.gov

The biogeochemical functioning of ecosystems is shaped in part by genomic information stored in the subsurface microbiome. Cultivation-independent approaches allow us to extract this information through reconstruction of thousands of genomes from a microbial population. Analysis of these genomes, in turn, gives an indication of the organisms present and their functional roles. However, metagenomic analyses can currently deliver thousands of different genomes that range in abundance/importance, requiring the identification and assimilation of key physiologies and metabolism to be represented as traits for successful simulation of subsurface processes.

Here we focus on incorporating -omics information into BioCrunch, a genome-informed trait-based model that represents the diversity of microbial functional processes within a reactive transport framework. This approach models the rate of nutrient uptake and the thermodynamics of coupled electron donors and acceptors for a range of microbial metabolisms including heterotrophs and chemolithoautotrophs. Metabolism of exogenous substrates fuels catabolic and anabolic processes, with the proportion of energy used for cellular maintenance, respiration, biomass development, and enzyme production based upon dynamic intracellular and environmental conditions. This internal resource partitioning represents a trade-off against biomass formation and results in microbial population emergence across a fitness landscape.

Biocrunch was used here in simulations that included organisms and metabolic pathways derived from a dataset of ~1200 non-redundant genomes reflecting a microbial community in a floodplain aquifer. Metagenomic data was directly used to parameterize trait values related to growth and to identify trait linkages associated with respiration, fermentation, and key enzymatic functions such as plant polymer degradation. Simulations spanned a range of metabolic complexities and highlight benefits originating from simulations including a larger number of organisms that more appropriately reflect the *in situ* microbial community. Simulations demonstrate the importance of dynamic vs. static energy allocation to predict active processes such as anammox. In addition, even though traditional methods showed no thermodynamic limitations on reaction rates, the maximum potential growth rate of iron reducers were shown to vary spatially by a factor of 2-3 due to differences in energy allocation.

This research is part of the Berkeley Lab Subsurface Biogeochemical Research Genomes-to-Watershed SFA.